Evaluating the Performance and Progress of Learning-enabled Systems

NOTE: The Solicitations and topics listed on this site are copies from the various SBIR agency solicitations and are not necessarily the latest and most up-to-date. For this reason, you should use the agency link listed below which will take you directly to the appropriate agency server where you can read the official version of this solicitation and download the appropriate forms and rules.

The official link for this solicitation is: http://www.acq.osd.mil/osbp/sbir/solicitations/index.shtml

Agency:
Department of Defense

Release Date:
November 20, 2013
Branch:
Office of the Secretary of Defense

Open Date:
November 20, 2013
Program / Phase / Year:
SBIR / Phase I / 2014

Application Due Date:
January 22, 2014

Solicitation:
 2014.1

Close Date:
January 22, 2014
Topic Number:
OSD14.1-AU2

Description:
This topic is supported under National Robotics Initiatives (NRI). OBJECTIVE: Develop methodology to evaluate and measure the performance and progress for learning enabled systems. DESCRIPTION: A long term goal of machine learning is to develop systems that learn complex behaviors with minimal human oversight. However, future systems that incorporate learning strategies will not necessarily have a fixed software state that can be evaluated by the testing community. In some cases, most of the training occurs in the development process using large databases of training examples. Testing may involve a series of challenge scenarios, similar to the DARPA autonomous mobility challenges, designed to examine the performance of the system-under-test in relevant conditions. Design of the scenarios and performance metrics are open research questions. As autonomous systems are used in increasingly complex scenarios, supervised training during the development phase, by itself, may not be sufficient to allow the system to learn the appropriate behavior. Learning from demonstration uses examples, often supplied by the end-user of the system, to train the system. Examples include flying in specific environments, bi-pedal locomotion on different terrain surfaces and, throwing objects of different sizes or densities. In this case, the tester needs to stand in for the end user and"train"the systems before testing it. Test procedures need to evaluate not only the performance of the system in various scenarios, but the amount of time it takes to train the system and the required level of expertise for the"expert"trainer. Finally, some applications include continuously adapting models that adjust over time to compensate for changes in the environment or mission. Current research is exploring the use of on-line learning in areas such

as terrain adaptive mobility and perception. This case presents a particularly challenging evaluation problem, performance in a given scenario is not static it may improve over time. In this solicitation we seek a methodology that answers the following three questions: a) What is an appropriate testing methodology for learning-enabled systems? This includes testing procedures that apply to systems with supervised learning components, as well as user-trained or continuously adapting systems. b) Are there general testing principles that can be applied to learning-enabled systems regardless of the specific applications? c) Can we predict the evolution of a learning-enabled system over time? For adaptive systems, can we predict how much time is required to adapt to a new environment? What are the potential impacts on military autonomous systems? PHASE I: The first phase consists of initial methodology development, metrics and a set of use cases to evaluate and measure the performance of learning-enabled systems. This methodology must address supervised, re-trained and continuously adaptive systems. Documentation of the methodology and use cases is required in the final report. PHASE II: Prototype the methodology by using it to examine test cases for each type of learning-enabled system in simulated test environments. The prototypes should address the 3 questions states above. Deliverables shall include the prototype system and a final report, which shall contain documentation of all activities in the project and a user's guide and technical specifications for the prototype system. PHASE III: Fully developed systems that evaluate the performance learning-enabled systems in either real or simulated scenarios. Potential commercial applications could be a system that to assess the performance of autonomous driving systems, logistics systems, and autonomous UAV applications such as power line inspection in which the UAV must adapt its flight parameters to changing wind characteristics. Deliverables shall include the methodology, test case scenarios and some general principles that the test and evaluation community can use to develop test procedures for specific systems.